# Draft genome sequence of Rhodococcus sp. strain NCIMB 12038, a naphthalene-degrading bacterium

AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements™

PROKARYOTES

# Draft Genome Sequence of *Rhodococcus* sp. Strain NCIMB 12038, a Naphthalene-Degrading Bacterium

**Shamsudeen U. Dandare,[a] Timofey Skvortsov,[a] Ksenia Arkhipova,[a]* Christopher C. R. Allen[a]**

[a]School of Biological Sciences, Queen's University Belfast, Belfast, United Kingdom

**ABSTRACT** We report here the draft genome sequence of *Rhodococcus* sp. strain NCIMB 12038, an industrially important bacterium, possessing a large and diverse repertoire of genes involved in the biotransformation of various organic compounds, including naphthalene.

The members of the genus *Rhodococcus* are genetically and metabolically diverse, which allows them to adapt to a wide range of environmental conditions and utilize various organic compounds as energy and nutrient sources, including highly toxic ones. *Rhodococcus* sp. strain NCIMB 12038, originally isolated by Larkin and Day from garden soil in 1983, accepts carbaryl as a sole carbon and nitrogen source (1, 2). The ability of this strain to metabolize and degrade complex organic molecules, including naphthalene, and produce valuable activated aromatic compounds, as well as its involvement in fossil fuel biodesulfurization, makes it an industrially and ecologically important microorganism (3, 4).

*Rhodococcus* sp. strain NCIMB 12038 was revived from a freeze-dried sample (prepared ca. 1998 in our laboratory) by plating on LB agar; a single colony was then picked and grown on solid minimal salt medium (MSM) with the addition of naphthalene as the only carbon source. Subsequently, a single colony was picked and grown in liquid MSM with the direct addition of 2 g/liter of naphthalene to the medium. In all growth experiments, bacteria were grown at 25°C to late exponential phase.

Cell cultures were pelleted, and genomic DNA was extracted using the FastDNA SPIN kit for soil (MPBio, Solon, OH, USA) following the manufacturer's protocol. Whole-genome sequencing was performed at the MR DNA sequencing facility (Shallowater, TX, USA) using the Illumina MiSeq platform. The paired-end 150-bp sequence reads generated were assembled at MR DNA using NGen DNA assembly software (DNAStar, Inc., Madison, WI, USA), producing 112 contigs, with an average coverage of 40×. Minimus2 (5) was used to merge contigs with overlaps, and the resulting set of 109 contigs was analyzed with CheckM (6), which characterized the draft genome as 99.5% complete. The assembly was annotated online using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (7).

The draft genome of *Rhodococcus* sp. NCIMB 12038 is 9.3 Mb long and has a GC content of 67.2%, which is in the range of values characteristic to other members of the genus (8). A total of 8,506 genes were predicted by NCBI PGAP, including 8,121 protein-coding sequences, 321 pseudogenes, and 64 RNA genes. Among the identified RNA genes, 49 were annotated as tRNAs and 3 as noncoding RNAs, while 4 copies of 5S, 3 copies of 16S, and 5 copies of 23S rRNAs were detected in the draft genome. SpecI (9) prediction and comparative analysis of the 16S rRNA sequences identified *R. opacus* as the closest related species. Annotation of the protein-coding genes with RASTtk (10) assigned 26% of them to functional categories of SEED subsystems, while analysis with AntiSmash version 4.0 (11) predicted 160 biosynthetic clusters. IslandViewer (12) predicted 76 genomic islands; several putative insertion sequences and prophages were

**Received** 13 November 2017 **Accepted** 14 November 2017 **Published** 4 January 2018

**Citation** Dandare SU, Skvortsov T, Arkhipova K, Allen CCR. 2018. Draft genome sequence of *Rhodococcus* sp. strain NCIMB 12038, a naphthalene-degrading bacterium. Genome Announc 6:e01420-17. https://doi.org/10.1128/genomeA.01420-17.



Address correspondence to Christopher C. R. Allen, c.allen@qub.ac.uk.

* Present address: Ksenia Arkhipova, Theoretical Biology and Bioinformatics, Utrecht University, Utrecht, the Netherlands.

S.U.D. and T.S. contributed equally to this work.

also detected. No clustered regularly interspaced short palindromic repeat (CRISPR) arrays were found after an inspection with CRISPRDetect (13).

**Accession number(s).** This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number NHML00000000. The version described in this paper is the first version, NHML01000000.

## ACKNOWLEDGMENTS

This work was partly funded by the Agri-Food Quest Competence Centre (AFQCC) at Queen's University Belfast, United Kingdom, and by a Commonwealth Doctoral Scholarship granted to S.U.D.

## REFERENCES

1. Larkin MJ, Day MJ. 1985. The effect of pH on the selection of carbaryl-degrading bacteria from garden soil. J Appl Bacteriol 58:175–185. https://doi.org/10.1111/j.1365-2672.1985.tb01445.x.
2. Larkin MJ, Day MJ. 1986. The metabolism of carbaryl by three bacterial isolates, *Pseudomonas* spp. (NCIB 12042 & 12043) and *Rhodococcus* sp. (NCIB 12038) from garden soil. J Appl Bacteriol 60:233–242. https://doi.org/10.1111/j.1365-2672.1986.tb01078.x.
3. Boyd C, Larkin MJ, Reid KA, Sharma ND, Wilson K. 1997. Metabolism of naphthalene, 1-naphthol, indene, and indole by *Rhodococcus* sp. strain NCIMB 12038. Appl Environ Microbiol 63:151–155.
4. Larkin MJ, Allen CC, Kulakov LA, Lipscomb DA. 1999. Purification and characterization of a novel naphthalene dioxygenase from *Rhodococcus* sp. strain NCIMB12038. J Bacteriol 181:6200–6204.
5. Treangen TJ, Sommer DD, Angly FE, Koren S, Pop M. 2011. Next generation sequence assembly with AMOS. Curr Protoc Bioinform Chapter 11:Unit 11.8. https://doi.org/10.1002/0471250953.bi1108s33.
6. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055. https://doi.org/10.1101/gr.186072.114.
7. Tatusova T, Dicuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res 44:6614–6624. https://doi.org/10.1093/nar/gkw569.
8. Pathak A, Chauhan A, Blom J, Indest KJ, Jung CM, Stothard P, Bera G, Green SJ, Ogram A. 2016. Comparative genomics and metabolic analysis reveals peculiar characteristics of *Rhodococcus opacus* strain M213 particularly for naphthalene degradation. PLoS One 11:e0161032. https://doi.org/10.1371/journal.pone.0161032.
9. Mende DR, Sunagawa S, Zeller G, Bork P. 2013. Accurate and universal delineation of prokaryotic species. Nat Methods 10:881–884. https://doi.org/10.1038/nmeth.2575.
10. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep 5:8365. https://doi.org/10.1038/srep08365.
11. Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, Lee SY, Fischbach MA, Müller R, Wohlleben W, Breitling R, Takano E, Medema MH. 2015. AntiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res 43:W237–W243. https://doi.org/10.1093/nar/gkv437.
12. Dhillon BK, Laird MR, Shay JA, Winsor GL, Lo R, Nizam F, Pereira SK, Waglechner N, McArthur AG, Langille MGI, Brinkman FSL. 2015. IslandViewer 3: more flexible, interactive genomic island discovery, visualization and analysis. Nucleic Acids Res 43:W104–W108. https://doi.org/10.1093/nar/gkv401.
13. Biswas A, Staals RHJ, Morales SE, Fineran PC, Brown CM. 2016. CRISPRDetect: a flexible algorithm to define CRISPR arrays. BMC Genomics 17:356. https://doi.org/10.1186/s12864-016-2627-0.